

~~CONFIDENTIAL~~
Security Information

APPENDIX IV

Machines for Searching Recorded Information

Automatic Classifying System developed by
Kybernetes Corporation

(A Video Approach)

I. Introduction

a. In March, 1953, the Kybernetes Corporation (a division of the Self Winding Clock Company, 9 East 40th Street, New York, N. Y.) demonstrated for the first time a working model of an "automatic classifying system" for the special purpose of producing, controlling, and utilizing magazine subscription records. A major component of the system is a set of equipment which searches recorded information in accordance with desired criteria. On the basis of the information found several independent functions, including reproduction, are performed selectively and automatically. The equipment utilizes a closed-circuit industrial television system.

b. In view of the apparent potentialities of the video approach it was thought desirable to investigate the Kybernetes system from the point of view of its applicability to general information searching. A preliminary investigation has been made, the results of which are outlined below. In summary, the following tentative conclusions were reached: The Kybernetes system is in principle applicable to general information searching, but details of the present design limit its usefulness for this purpose. By straightforward engineering modifications it should be possible to achieve a revised design which would be highly effective for searching purposes.

II. Brief Description of the Kybernetes System

a. Production of Records.

The Kybernetes 'automatic classifying system' was designed for the special purpose of producing and handling magazine

Security Information

subscription records. For the production of needed records a master-slave typewriter assembly has been provided which features automatic coding; the typical product of this assembly is a coded file card and an uncoded bill; name and address of subscriber, etc., are typed manually on the file card and slave-typed on the bill. Those portions of the typed information which are to be searched mechanically are typed automatically on the file card in mechanically scannable code; each code "character" consists of a linear pattern of one or more black bar-shaped spots, typed spot by spot in a horizontal four-place field. The file cards now employed are white cards of standard tabulating size. One card is prepared for each subscriber.

b. Searching System and Output Functions

For the control and utilization of card files generated in the above manner, a set of searching equipment has been developed which is capable of executing a variety of functions on a single pass of cards. The basis of this equipment is a closed-circuit industrial television system. In a typical operation of the searching system a file of coded record cards is fed and the coded information is scanned by a television camera; on the basis of the scanned information obsolete record cards (expired subscriptions) are sorted out; needed counts are taken (for example, counts for bulk mailing purposes); and through the action of a second television camera, several different selective listings are generated (for the provision of mailing labels and for other purposes). A listing consists in a photographic reproduction (partial or total) of the information-bearing side of those record cards whose coded information meets the requirements for selection. A given record card is reproduced simultaneously on several listings if it happens to meet their respective requirements. The various requirements for listing, sorting, and counting are presented to the searching equipment by the setting of switches.

III. Features of the Present Design which Restrict its Usefulness for Information Searching

If the 'automatic classifying system' were applied in its present form to information searching, the following features would tend to limit its usefulness:

- a. The information susceptible to searching is now recorded and searched in fixed fields.

~~CONFIDENTIAL~~
Security Information

- b. The 'alphabet' used for recording searchable information now comprises only ten (10) patterns (corresponding to ten numeric characters).
- c. The density at which searchable information is now recorded corresponds to about 390 spaces or bar-positions per unit record of standard tabulating size. (Compare this with 960 hole-positions per IBM card.)
- d. The rate at which unit records (standard tabulating size) pass through the searching equipment is relatively slow:
 - ca. 240 cards per minute, according to design;
 - ca. 120 cards per minute, as demonstrated.
- e. The scope and flexibility of searching operations for which provisions have been made are limited to the moderate requirements of magazine subscription control.

The ability to search for combinations of criteria exists in the present design; there has been no occasion as yet, however, to incorporate a holdover function or clearing functions, or to provide conveniently for searching on the basis of a relatively large number of diverse criteria, variable in length, and combinable in a variety of ways.

IV. Modification of Present Design to Improve Applicability to General Information Searching

The limiting features indicated above were discussed with representatives of the Kybernetes Corporation, and it was found that none of them is basic to the system. By straightforward engineering changes it should be possible to alter the design in the following respects, each of which would improve the design for the purpose of general information searching:

- a. Provisions could be made for recording and searching information without recourse to fixed fields.

The typewriter assembly for producing records would be re-designed, and it seems probable that a flat-bed design would be adopted.

- b. The 'alphabet' of searchable patterns could be expanded as needed.

For example, an alphabet of six-place patterns could be employed which would accommodate both alphabetic and numeric characters.

c. The density at which searchable information is recorded could be increased substantially.

(1) Without reducing bar size, the density could be increased to about 1500 bar-positions per unit record of standard tabulating size. (Compare this figure with 960 hole-positions per IBM card.)

This would be done by changing from double spacing to single spacing between (a) bars and (b) lines.

(2) By reducing bar size and spacing, further increases in density could be realized.

It is not yet known how much increase of this sort is immediately practicable; it may become feasible to go to a reduction which would permit a typewritten character together with its bar-pattern to be incorporated on a single small type bar.

d. The effective rate at which unit records are fed and scanned could be increased substantially, in one or more ways.

(1) The absolute rate probably could be increased to at least 600 unit records per minute.

Integration of the two camera stations into a single station would be helpful in this respect, as well as in others.

(2) The designed rate could be doubled, in effect, by synchronizing two feeding units with a single electronic-video system.

e. The scope and flexibility of searching operations could be increased as needed, by the provision of appropriate circuitry and keyboards for its manipulation.

V. Effectiveness of Modified Design for General Information Searching

If changes of the foregoing types were made, the resulting design should compare favorably with the most advanced punched-card systems thus far proposed (for example, the IBM Electronic Scanning System, Mark III). It appears that the modified design would have also the following distinctive features:

- a. The dimensions of the unit record could be fixed at an optimum size.

For example, a 5" by 8" card might be preferred to a card of standard tabulating size. A relatively "large" or "long" card could be considered if arrangements were made to permit the recording and searching of more than one source of information per card.

- b. The possibility of increasing the density of searchable information could be rather freely exploited, since the choice of bar size and spacing would not be deterred by a heavy investment in existing equipment (key punches, etc.).

- c. It would be possible to reproduce information from both sides of cards. (The searching of both sides also could be considered.)

- d. Since reproduction is a major output function,

- (1) it would be possible to make the system one in which several complete searches (for example, five or more) could be conducted on a single pass of cards.

This might be an important factor from the point of view of overall speed and efficiency of operation and average cost per search.

- (2) the problem of controlling the completeness and sequence of sets of cards (i.e., the follower-card problem) could be minimized.

- e. The choice between unit records and a continuous medium would be open to determination; either or both types of feeding unit could be provided; and in each case, the reproductions generated could be provided in a form amenable to further searching (as in the case of punched cards physically selected).

VI. Miscellaneous Notes

a. Note Concerning Reproduction

For the production of listings, the system now utilizes rolls of fast photographic paper; this paper is imported and is eight to ten times faster than domestic papers now available. The exposed paper is processed on a continuous basis in

~~CONFIDENTIAL~~
Security Information

standard commercial equipment. To reduce curling the use of softening solutions is under investigation. (Question: Is an appropriate dry process now available?) A representative of Kybernetes suggested that for certain applications it might be desirable to substitute for photographic reproduction a high-speed automatic typewriter of the Analex type. (Analex, made in Boston, operates at a rate of 600 characters per minute.)

b. Note Concerning Cost of System

The working model of the 'automatic classifying system' is not intended to be a prototype for large scale production. Instead, Kybernetes Corporation has in mind that any equipment built hereafter will be tailored to the user's needs. They would be willing to build either for rental or for purchase. The cost of rental or purchase would be determined on the basis of the user's requirements. If the system were modified for application to general information searching, it would be helpful if the cost of re-design could be distributed initially among several users (for example, several industrial firms or government agencies).

VII. Conclusions

Preliminary investigation of the Kybernetes 'automatic classifying system' has indicated that the system is applicable in principle to general information searching and that a system suitable for general searching could be achieved by straightforward modification of the present design. In view of this adaptability, together with the distinctive features of the system, it is in order to conclude that further attention should be given to the video approach to searching; in particular, it would be helpful if arrangements could be made to obtain from Kybernetes Corporation an estimate of cost and building time for a set of appropriate equipment.

* * *

- 6 -

~~CONFIDENTIAL~~
Security Information